AFGHAN REFUGEE CAMP SURVEYS: PAKISTAN, 2002

James Bell, David Nolle, Ruth Citrin, U.S. Department of State
and Fritz Scheuren, NORC, University of Chicago

Key Words: Area Probability Design, Design Effects, Survey Contract Management, Acquiescence Bias in Interviews

Abstract: Both as professionals and as citizens, the events of the last year and now more have brought about many changes to our view of the world and our engagement in it. This survey was one response to those changes. Its main goal was measuring attitudes on a variety of social, economic, and political issues of the Afghan refugees that are now returning to their homeland from Pakistan. Particularly important was learning about their perceptions regarding current circumstances as well as future expectations. From a methodological perspective, trying to obtain a good sample of adult males in the refugee camps posed many challenges and most of the discussion will be focused on these challenges.

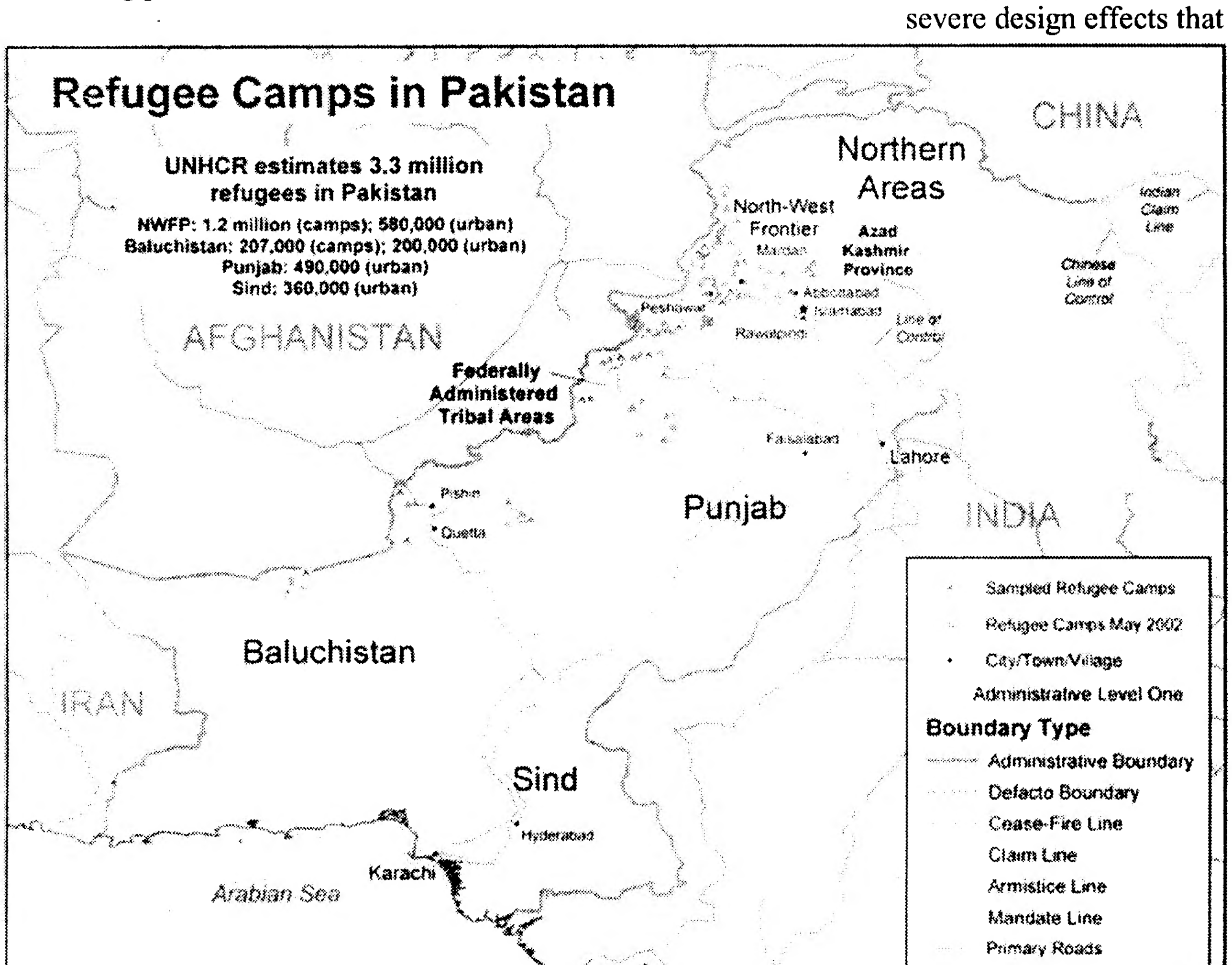
1. Introduction

This paper provides initial findings from a survey conducted this past spring in a sample of Afghan refugee camps in Pakistan.

The survey is an unusual but very important and timely example from the extensive overseas research program that has been mounted historically by the U.S. government, initially by the U.S. Information Agency and now at the Office of Research, Department of State. That Office is the official pollster for the U. S. Government abroad and is involved in conducting over 120 polls in over 70 countries each year. These polls are policy-driven and typically cover a wide variety of topics, including political and economic trends within a given country, perceptions of the U.S., bilateral security, and views of democracy.

As you might guess, the tragic events of September 11, 2001 have added to our topics of concern. Although most of our surveys involve urban or national probability designs focused on adults in a given country, the survey being discussed today is an important exception because we are focusing on a very different kind of population.

To orient the reader we have provided a map (see below) showing the spatial distribution of Afghan refugee camps within Pakistan. Notice there is a high concentration of camps in the North West Frontier Province (NWFP) bordering eastern Afghanistan.



From the standpoint of its organization, the paper is divided into seven brief sections, including this introduction (Section 1). In Section 2 the main purposes of the research and the consequent analytic focus are detailed. Next (in Section 3) the sample design is described, including how we used the camp lists provided by the United Nations High Commissioner for Refugees (UNHCR) and the limitations imposed by such a frame. The rapid development we had to do of the questionnaire in two languages, Pashto and Dari, is covered in Section 4, along with some initial results. Because this population does not have a high literacy rate and is generally unfamiliar with polling procedures, our questions were greatly simplified; consequently, some of the results from this survey probably reflect some acquiescence bias (see, for example, Javeline, 1999). Section 5 describes how we handled various contract survey management issues, some successfully, some less so. Information on the severe design effects that we encountered is presented in Section 6. Further analyses are touched on in the final section, Section 7, entitled "Next Steps."

2. Research Purposes and Consequent Analytic Focus.

Because Afghan refugees will eventually take part in deciding Afghanistan's political and economic future, we wanted, in the survey discussed here, to gauge Afghan refugee attitudes on a variety of social, economic, and political issues. In addition, because we expected that their current circumstances as well as their future expectations might influence their desires to repatriate to Afghanistan, we wanted to get perceptions of their current situations, as well as their expectations regarding their futures in Afghanistan.

These purposes pushed us in the direction of a comprehensive attitudinal survey (See Section 4). The design called for coverage of Afghan refugee males as well as

PROPERTY

females -- located in camps as well as in urban areas throughout Pakistan. However, because we anticipated that, for social and cultural reasons, males would play the dominant role in any decision to repatriate, we focused a major portion of our resources on designing the survey to get a good sample of males in the refugee camps. Only the male refugee sample drawn in the refugee camps is described in what follows.

3. Sample Design and UNHCR Frame

To sample male, mainly household heads, in the refugee camps we began with the camp lists provided by the United Nations High Commissioner for Refugees (UNHCR). This frame was dated and, as we expected somewhat incomplete. Even so, the UNHCR camp counts at the time of sampling were believed to offer roughly accurate size measures of the total residency in each camp and were relied on for sampling.

Our design was quite standard in most respects but its implementation had many challenges. For frame construction, we had two basic lists: A lengthy list of names and population estimates for 92 older camps, which had been established during various time periods prior to the October 7, 2001 bombing of the Taliban in Afghanistan. There was also a shorter list of 8 new camps that had emerged to handle the refugees from the October 7th bombings. In both cases, we wanted to use camps as the primary sampling unit (PSU).

Although we recognized that there were problems with the list of older camps, we used this list as the basis for all of our subsequent sampling activities for camps established prior to the October 7th bombings. Consequently, we shall focus on the sampling details related to these two lists for the bulk of the presentation.

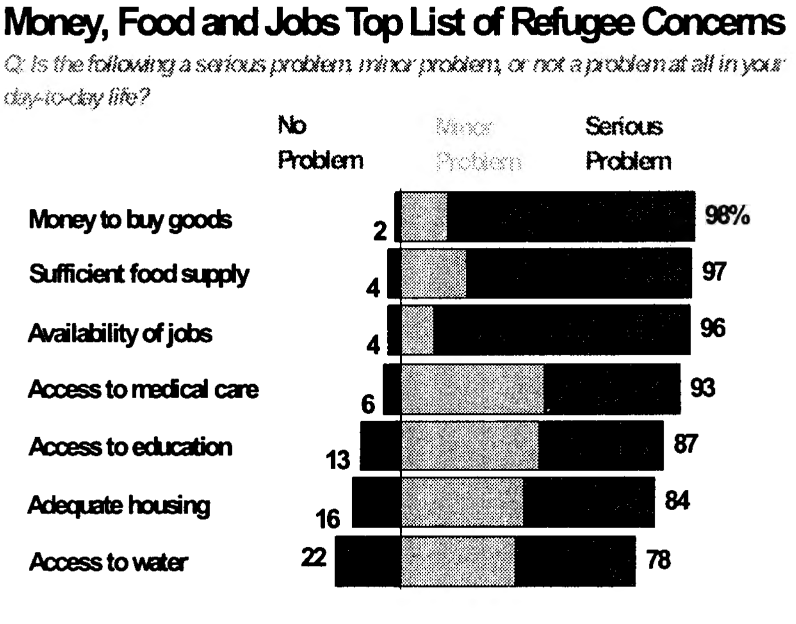
At the outset, we allocated 1200 interviews to the old camps and 200 interviews to the new camps. We anticipated that post-stratification weighting would rectify any disproportionality that we introduced with this initial allocation and our remaining stratification. Because the old camps varied in size from approximately 4,000 to over 75,000, we stratified them into nine relatively homogeneous groups based on rounded estimates of their population sizes as well as their locations. For example, at one end of the continuum in NWFP, our first stratum contained 24 camps averaging around 10,000 refugees per camp while our second stratum contained 27 camps averaging around 15,000 refugees per camp. At the other end of the continuum, we had three very large camps ranging from 55,000 to 75,000 persons, and we selected each of these with certainty.

Within each homogeneous stratum having multiple camps, we selected the camps with equal probability, and we allocated 40 interviews to each randomly selected camp. We felt that this approach would give us a relatively equal probability of inclusion for all respondents selected within a given stratum. Furthermore, we chose the number of camps to correspond approximately to the estimated size of the population in the stratum to avoid any major disproportionalities among strata. In addition, within each stratum, we divided the randomly selected camps into two random halves so that in the end we had 12 camps going into one random half and 12 camps going into the other random half. We also subdivided each of the three largest camps into random halves where we put 40 interviews into each half, so that we ultimately had two groups of camps each having a total of 15 random halves. We felt that the use of random halves would provide us not only with insurance in the event that the project was disrupted by rapid repatriation or significant shifts in the refugee population but also with reassurance when we analyzed these random halves to see whether they were consistent or divergent in any meaningful ways. Furthermore, the random halves for the camps selected with certainty provided us with a basis for our estimation of the within camp sampling variance in the certainty cases.

4. Questionnaire Design and Some Initial Results

Most of the questions in the questionnaire had to be restricted to reasonably simple forms. For example, five point Likert scales were replaced by simple agree-disagree dichotomies. The amount of work to construct even this type of questionnaire was considerable. The questions had to be in two languages, Dari and Pashto. As a part of our procedures, both the Dari and the Pashto versions of the questionnaire were independently back-translated by Dari and Pashto translators in the U.S. and changes were forwarded to the contractor by faxes.

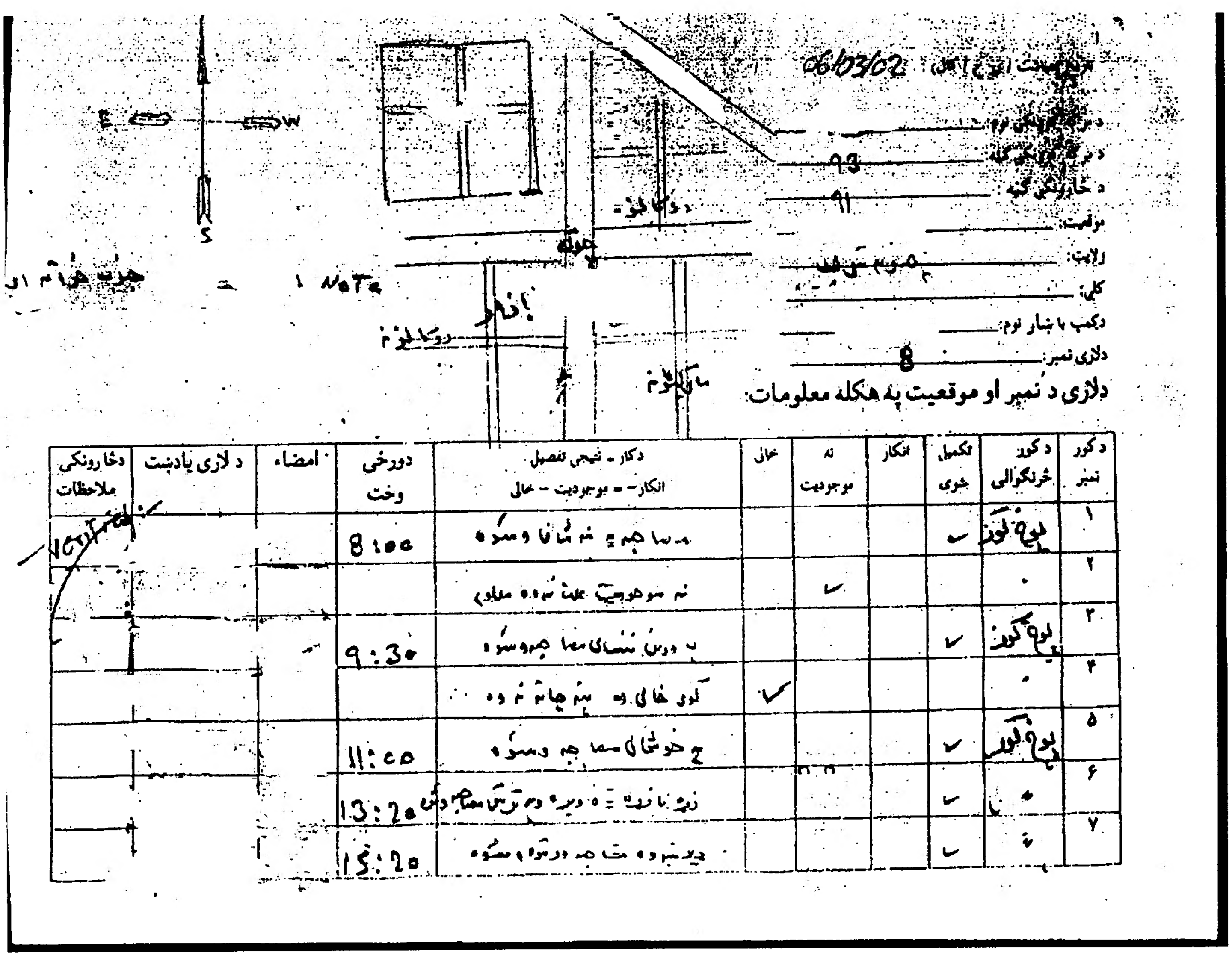
Judging from some earlier related work done by one of our colleagues (Javeline, 1999), we knew that there was likely to be some acquiescence bias or "agreement bias" using the agree-disagree approach – that is, that there would be an overestimation of agreement with an opinion, say **Opinion A**, when the question is phrased "Do you agree or disagree with **Opinion A**?" Our preliminary analyses suggest that at least some politically-oriented questions may have been subject to such bias. However, many of our other questions on social and economic issues and living conditions have different types of question formatting and response alternatives. Thus, it is important to emphasize that the relative ranking of the results shown in Table 1 at the end of this paper are probably unaffected by our concern about acquiescence bias. Moreover, many opinions were so nearly universal that we feel comfortable in treating them as reliable enough to act on. For example, the bar graph shown below highlights the respondents' assessments of their living conditions. Given some of the dire circumstances that are evident in these camps, it is no surprise that access to money, access to food, and the unavailability of jobs top the list of serious problems. It is very significant that even access to a basic such as water is a serious problem for many respondents. The full texts for this series of questions as well as other questions used in our initial analyses are reported in Table 1, which also includes the weighted percentages as well as the jackknifed standard errors for these percentages.



5. Survey Management

For this survey, most of the time the in-country contractor was half a world away from the principal investigators. In fact, the amount of direct onsite supervision was limited to just one face-to-face set of meetings. Given this setting we had to introduce a number of (remote) control procedures to prevent, if possible, and detect, when necessary, procedural deviations. And we did find some deviations, but they were ***not*** judged to bias the results to any serious extent, although they did contribute to an increase in the measured sampling error, about which more will be said in Section 6. What then did we do? Some examples will have to suffice.

For the pilot survey, the contractor spent one day at a major camp in an effort to pre-test the questionnaire. We also wanted the contractor to check on the usability of the random route approach we were using to get random selections in each camp (see Holly, 2002 for related ideas) and to check their skip intervals and turning procedures as they moved along their routes (see the Pashto version of a completed random route contact sheet below; note that unique identifying information has been removed). As a result of the pilot, both the questionnaire and the random route contact sheet were revised.



Although the start points and walking directions were not picked by a completely random procedure, our feeling is that any deleterious impacts from this decision may have been minimized by the fact that eight random routes were going to be assigned in each camp and only five or so interviews were going to be conducted on each route. It was decided that any willing adult male, not necessarily the household head, would be interviewed in each selected dwelling unit.

There were a total of ten interviewing teams covering the males in the selected camps. Each interviewing team was composed of a supervisor and two interviewers, and all interviewers were Afghan refugee males. The allocation of the interviewing teams was supposed to follow a pattern so that no one team dominated a single stratum. In eight of the nine strata covering the old camps, this practice held true. Each of the twenty interviewers ultimately conducted between 40 and 120 interviews in this survey of 1400 males in 27 old camps and 5 new camps. The average interview was approximately 57 minutes in length. All of these interviews were completed between February 16 and March 22, 2002.

6. More Survey Results and Design Effects

The design effects that we encountered in this survey, as already noted, were unusually severe. These effects are presented graphically in the box and whisker plots shown below. This graphic presents the DEFT's (that is, the square roots of the design effects called DEFF's) for all of the response options relating to various types of items initially analyzed in our survey. These DEFT's have been computed from the DEFF's derived from Wesvar 4.0.

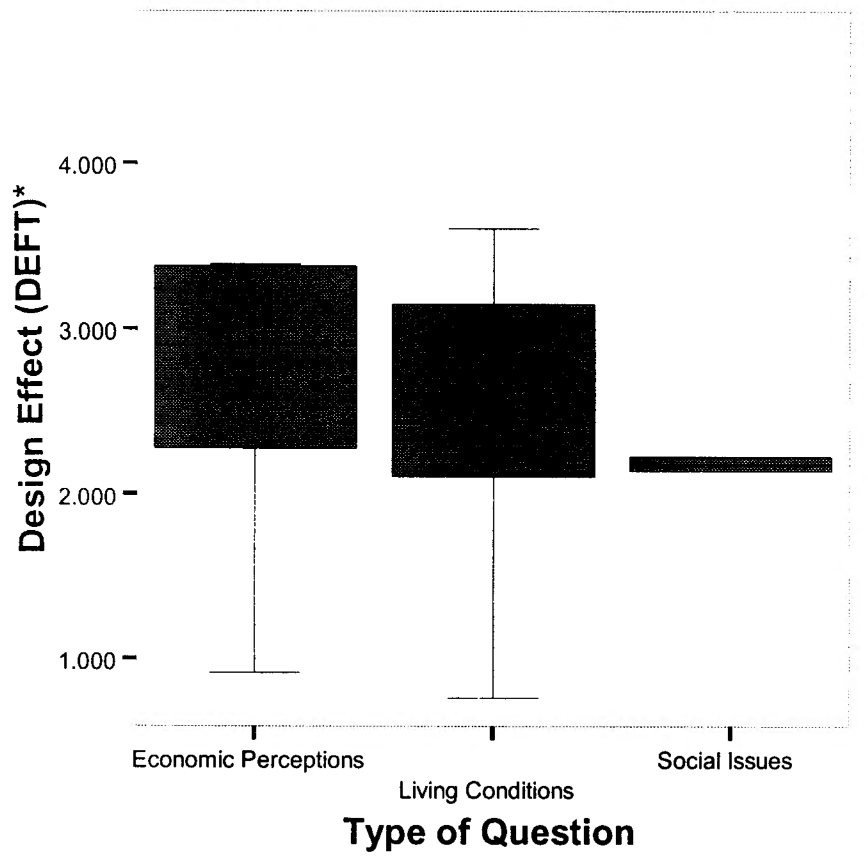
Each of the DEFT's incorporated into this graphic is a ratio of two kinds of standard errors: The numerator of this ratio is the standard error calculated by a replication method called "Jackknife n" while the denominator is the standard error computed with a simple random sample formula. In all of our preliminary calculations, the finite population correction factor has been ignored.

Through the use of "Jackknife n", our replicates for estimating the error variances in our sample are formed in a way which takes into account our stratification as well as the clustering which arises from the thirty-two camps used in our male sample. Then our replicate estimates are feed into a generalized formula that ultimately provides the standard error for any given statistic such as the percentage saying that a sufficient food supply is a serious problem. The core of this formula is simply the sum of the squared deviations of the replicate estimates of a given statistic (for example, the percentage saying that a sufficient food supply is a serious problem) from the full sample estimate of this statistic.

This chart summarizes the DEFT's for three types of items: economic perceptions; living conditions; and social issues. Three basic things are prominent in this chart:

Design Effects* By Question Type

(Afghan Males in Refugee Camps in Pakistan, 2002)



*A design effect is the ratio of the jackknifed standard error to the simple random sample standard error.

First, most of the DEFT's are larger than 2.0 which is considerably bigger than DEFT's found in typical national surveys around the world. As a matter of fact, no interquartile range (that is, the outer edges of the boxes) for any of the three types of questions falls below 2.0.

Second, with the exception of the DEFT's computed on selected social issues, these DEFT's vary considerably within each type of question.

Third and most importantly, these DEFT's vary substantially by type of question. For example, the median for the DEFT's relating to social issues is slightly above 2.0 while the median for the DEFT's on the three economic perceptions falls close to 3.0.

These findings have implications for reporting the results from this survey because it is unlikely that a single standard error computed for the total survey will realistically capture the sampling errors generated for these items. In other words, these items will probably warrant special consideration in error reporting.

Meanwhile, this sort of finding (involving large design effects) forces one to search for its sources. At the outset, we have eliminated case weighting as the key explanation for the large design effects in this chart. We have discovered that, although the weighted data predictably produce larger DEFT's than the unweighted data, the striking feature is that the increase is very modest: 6.6% in the case of the mean DEFT and 3.1% in the case of the median DEFT.

Given our analysis of the demographics of some of these camps, we have come to believe that certain common backgrounds, shared experiences, and possibilities for long-term and frequent social interactions have created strong attitudinal clustering, which is likely to be the primary source of these particularly large design effects. In other words, it seems likely to us that most of the design effects for these items appear to be the byproducts of clustering in this sample.

7. Next Steps

We are continuing our analysis in several ways. For example, we are reviewing the effects of the presence of others on the interviewee's responses. The table below suggests that these effects are modest.

Detailed analyses of item nonresponse rates indicate that the item nonresponse rates were relatively low and comparable between the random halves formed during the initial sampling of the camps.

Meanwhile, further analyses are needed to determine whether the relatively large design effects found for our initial set of items reported herein extend to other types of questions such as ones on social backgrounds and open-ended questions.

			WERE OTHERS PRESENT WHEN THE INTERVIEW WAS CONDUCTED		DID ANYONE OTHER THAN THE RESPONDENT LISTEN TO THE QUESTIONS ASKED?		WERE QUESTIONS ANSWERED FREELY BY THE RESPONDENT?	
		Total	1 YES	2 NO	1 YES	2 NO	1 YES	2 NO
DESCRIBE YOUR FAMILY'S ECONOMIC CIRCUMSTANCES THESE DAYS	1 GOOD	3.0%	6.8%	2.7%	7.7%	2.6%	2.7%	7.4%
	2 OK	38.2%	48.3%	37.5%	47.7%	37.5%	38.1%	39.6%
	3 BAD	58.3%	44.8%	59.3%	44.6%	59.3%	58.6%	53.0%
	8 NO ANSWER	.1%	.0%	.1%	.0%	.1%	.1%	.0%
	9 DON'T KNOW	.4%	.0%	.4%	.0%	.4%	.4%	.0%
	Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	Weighted Sample	1400	95	1305	94	1306	1330	70
DO YOU EXPECT YOUR FAMILY'S ECONOMIC CIRCUMSTANCES WILL IMPROVE 1-2 YEARS FROM NOW	1 IMPROVE	15.5%	34.4%	14.1%	36.2%	14.0%	15.0%	25.1%
	2 REMAIN SAME	17.9%	10.1%	18.4%	9.2%	18.5%	18.2%	11.7%
	3 WORSEN	36.3%	25.6%	37.1%	23.7%	37.2%	36.7%	30.0%
	8 NO RESPONSE	1.1%	1.4%	1.1%	1.4%	1.1%	1.0%	3.7%
	9 DON'T KNOW	29.2%	28.4%	29.3%	29.4%	29.2%	29.2%	29.6%
	Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	Weighted Sample	1400	95	1305	94	1306	1330	70
DO YOU FEEL YOUR FAMILY IS BETTER OFF THAN WHEN YOU LIVED IN AFGHANISTAN	1 BETTER OFF	11.2%	17.0%	10.8%	17.4%	10.8%	10.7%	21.7%
	2 ABOUT SAME	8.0%	12.9%	7.7%	12.3%	7.7%	8.0%	7.7%
	3 WORSE OFF	79.6%	67.8%	80.4%	69.0%	80.4%	80.2%	68.8%
	8 NO RESPONSE	.1%	.0%	.1%	.0%	.1%	.1%	.0%
	9 DON'T KNOW	1.1%	2.3%	1.0%	1.3%	1.1%	1.0%	1.8%
	Total	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	Weighted Sample	1400	95	1305	94	1306	1330	70

Acknowledgements: The authors wish to acknowledge the very helpful feedback from the attendees at our presentation before a joint meeting of the Washington Statistical Society and the local chapter of the American Association of Public Opinion Research as well as excellent comments from Thomas Jabine.

REFERENCES

Javeline, Debra (1999). "Response Effects in Polite Cultures: A Test of Acquiescence in Kazakhstan," Public Opinion ***Quarterly, 63, 1-28.***

Reed, H. (2002). ***Demographic Assessment Techniques in Complex Humanitarian Emergencies***. National Academy Press.

Table 1: Selected Questions Used in the Analysis

(Note: Dari and Pashto versions of these questions were used in the field. The numbers in parentheses beside the response options are the weighted percentages for the options and their design-based standard errors derived from a jackknife replication method implemented within computer software called WesVar 4.0. An asterisk in parentheses indicates that no person gave this option. The analytical categorization of each question is noted in brackets after the stem of the question.)

Q2. How would you describe your own family's economic circumstances these days—is it good, "so-so", or bad? [Economic Perceptions]: Good (3.0%, 0.6%), OK (38.2%, 2.8%), Bad (58.3%, 3.0%), No answer (0.1%, 0.1%), Don't know (0.4%, 0.2%)

Q3. A year or two from now, do you expect your family's economic circumstances will improve, remain about the same, or worsen? [Economic Perceptions]: Improve (15.5% 3.1%), Remain about the same (17.9%, 2.6%), Worsen (36.3%, 3.7%), No response (1.1%, 0.7%), Don't know (29.2%, 4.1%)

Q4. All things considered, do you feel your family is better off, about the same, or worse off than you were when you lived in Afghanistan? [Economic Perceptions]: Better off (11.2%, 2.8%), About the same (8.0%, 1.3%), Worse off (79.6%, 3.6%), No response (0.1%, 0.1%), Don't know (1.1%, 0.3%)

Q5. I am now going to read you a list of issues related to the situation of Afghan refugees. For each one, please tell me if it is a serious problem, a minor problem, or no problem at all for your family? First take "access to water," is this a serious problem, a slight problem, or no problem at all for your family. Now take... [REPEAT FOR EACH ITEM ON LIST]

A. Access to water [Living Conditions]: Serious problem (40.4%, 6.3%), Minor problem (37.8%, 4.1%), No problem (21.8%, 3.8%), No response (*), Don't know (*)

B. Availability of jobs [Living Conditions]: Serious problem (84.6%, 1.7%), Minor problem (11.0%, 1.7%), No problem (4.4%, 0.9%), No response (*), Don't know (*)

C. Sufficient food supply [Living Conditions]: Serious problem (74.8%, 2.5%), Minor problem (21.6%, 2.2%), No problem (3.6%, 0.7%), No response (*), Don't know (0.0%, 0.0%)

D. Access to medical care [Living Conditions]: Serious problem (45.1%, 3.1%), Minor problem (47.8%, 3.3%), No problem (6.2%, 1.5%), No response (*), Don't know (0.9%, 0.5%)

E. Sufficient income to buy goods [Living Conditions]: Serious problem (81.6%, 2.2%), Minor problem (15.9%, 2.3%), No problem (2.0%, 0.4%), No response (0.1%, 0.1%), Don't know (0.4%, 0.2%)

F. Access to education [Living Conditions]: Serious problem (41.4%, 2.7%), Minor problem (45.5%, 3.1%), No problem (13.1%, 2.6%), No response (*), Don't know (0.1%, 0.1%)

G. Adequate housing [Living Conditions]: Serious problem (42.5%, 4.6%), Minor problem (40.9%, 4.3%), No problem (16.4%, 3.6%), No response (*), Don't know (0.3%, 0.1%)

Now, I'd like to ask you a few questions on some issues important to Afghans.
Q6. Thinking about education for Afghan children, do you believe that Afghan girls should or should not have the same educational opportunities as boys? [Social Issues]: Girls should have the same educational opportunities as boys (87.3%, 2.0%), Girls should not have the same educational opportunities as boys (11.5%, 1.5%), No response (0.2%, 0.1%), Don't know (1.0%, 0.6%)

Q7. What do you think about Afghan women working outside the home to help support their families? Do you think women should or should not be allowed to work outside the home? [Social Issues]: Women should be allowed to work outside the home (81.5%, 2.2%), Women should not be allowed to work outside the home (16.6%, 1.9%), No response (0.4%, 0.2%), Don't know (1.6%, 0.6%)

Q8. Sensitive question whose wording is not reported but whose percentages and standard errors are used in the analysis [Social Issues]: Combined (79.9%, 2.3%), Alone (19.3%, 2.3%), No response (0.1%, 0.1%), Don't know (0.7%, 0.3%)